

## **EVALUATING DRUG COURTS: A MODEL FOR PROCESS EVALUATION**

**By Cary Heck, Ph.D.,  
and Meridith H. Thanner, Ph.D.**

*Process evaluations are important tools for program management and oversight. Done well, drug court process evaluations should provide program managers with insight into their program's operations as they relate to the fundamental mission of improving the long-term prospects for their clients. Additionally, process evaluations of drug courts should promote consistent data collection and analysis of drug court activities. This article discusses and presents the elements and realities of process evaluation with the aim of assisting local programs in working with independent evaluators to develop and sustain ongoing process evaluation mechanisms. It is largely the product of the NDCI's National Research Advisory Committee and thus is the synthesis of suggestions from a broad base of well-known substance abuse and drug court researchers.*

*Cary Heck, Ph.D., is an Assistant Professor of Criminal Justice at the University of Wyoming as well as the Director of Research for the National Drug Court Institute. Previously, he served as Statewide Director of Drug Courts for Louisiana through the State Supreme Court. Dr. Heck has also worked as a small town police officer, a probation officer, a gang task force member, and a counselor for delinquent youth.*

*Meridith H. Thanner, Ph.D., is a Research Associate with the Bureau of Governmental Research (BGR) at the University of Maryland College Park and an independent consultant with the National Drug Court Institute. Most recently, her work with BGR has focused on managing process and performance evaluation studies of adult DUI courts and juvenile drug courts across the State of Maryland. Other research interests, stemming from her doctoral work in*

*military sociology at the University of Maryland, includes ethnographic assessments of how communities are affected by military base closings and realignments as part of the Federal Base Realignment and Closure Process.*

Direct all correspondence to Cary Heck, Ph.D., University of Wyoming, Criminal Justice, Department 3197, 1000 East University Avenue, Laramie, Wyoming 82071. (307) 766-2614; [check@ndci.org](mailto:check@ndci.org).

## ARTICLE SUMMARIES

### **WHAT ARE PROCESS EVALUATIONS?**

[8] Process evaluations are tools that drug courts should use to measure their efficiency, efficacy, and achievement of program goals.

### **WHO SHOULD CONDUCT THESE EVALUATIONS?**

[9] Trained evaluators have a skill set specific to the task and should be willing to listen to staff concerns.

### **WHAT ARE THE CRITICAL ELEMENTS OF PROCESS EVALUATION?**

[10] Critical elements of analysis include program goals, target population, drug treatment, court processes, units of service, team cooperation, and community support.

### **WHAT DATA ARE NEEDED TO COMPLETE THESE EVALUATIONS?**

[11] Data collected must be valid and reliable over time. Confidentiality guidelines must be followed in the collection of sensitive data.

### **WHAT DOES IT MEAN TO BE “METHODOLOGICALLY RIGOROUS”?**

[12] Even though experimental design may not be feasible, evaluation research must still follow accepted guidelines in its methodology.

### **WHAT ABOUT EXPERIMENTAL DESIGN AND COMPARISON GROUPS?**

[13] Quasi-experimental design is the most practical method of evaluating program outcomes and impacts. Comparison groups should be matched to drug court groups on the basis of research-established factors.

## INTRODUCTION

**D**rug courts are a national phenomenon. Few, if any, criminal justice interventions have spread throughout the country with as much speed and support as drug court programs. Drug court programs started with a single locally driven project in 1989 and have grown exponentially in number since. In 2004, the total number of operational drug court programs in the United States reached 1,621 and the total number of problem solving courts (e.g., mental health courts, domestic violence courts, family treatment courts) in general was 2,557 (Huddleston, Freeman-Wilson, Marlowe, & Roussell, 2005). The reason for this rapid growth is three-fold. First, drug court programs are based on the intuitive model of program design and implementation; that is, this person-centered model requires relatively little investment in order to realize long-term public support savings. Thus, the broad societal benefits of implementing a drug court program far out weigh the incremental costs involved in its formation and maintenance. Second, in light of the noticeable successes of drug court programs at the local level, a tremendous word of mouth public relations campaign has been undertaken by judges, drug court professionals, and clients, which has led to an impressive number of anecdotal cases in support of the model. Third, there is a growing body of empirical research that supports drug courts as effective programs for dealing with substance abusing offenders. Drug courts are credited with reducing recidivism, retaining clients in treatment, and improving outcomes and quality of life circumstances for substance abusing offenders. And, as evidenced in an increasing body of scientifically rigorous research, it is clear that drug courts are effective alternatives to traditional “business as usual” methods (Marlowe, DeMatteo, & Festinger, 2003).

## **The Current State of Local Drug Court Evaluation**

While the growth of drug court programs and the increasing body of literature suggesting the effectiveness of drug courts are undeniable, the quality and utility of many local program evaluations and data collection strategies remain questionable. Drug courts have historically faced considerable criticism in the area of evaluation and documentation. Through the Bureau of Justice Assistance (BJA), the federal government allocates millions of dollars to fund local drug court programs, but despite repeated efforts to count and document the activities of these programs, there is little uniform data on actual drug court activities nationwide. Congress has asked the U.S. Government Accountability Office (GAO) to review drug court evaluation and outcome research to determine the effectiveness of drug courts no less than four times. The most recent GAO Congressional Report on adult drug courts only found 27 of 117 evaluations of local drug court programs to be of sufficient methodologically quality to use for analysis (GAO, 2005). The findings from these evaluations indicate uniformly that drug courts produce positive results, but the lack of a broader selection of methodologically sound evaluations has led to continued skepticism.

In 2004, the National Drug Court Institute (NDCI) entered into an agreement with BJA and the National Institute of Justice (NIJ) to review evaluation plans for all federal drug court grantees. These evaluation plans were a special condition for grant recipients. NDCI reviewed over 100 evaluation plans and found considerable variation in the quality of the proposed research. Many of the evaluation plans required considerable technical assistance and a few were considered to be completely without merit and required rewriting. Programs often had difficulty identifying an appropriate and rigorous evaluation method, little theoretical understanding to guide their proposed evaluation activities, and poor data collection systems to support the conduct of a

methodologically sound process and performance evaluation. Furthermore, anecdotal evidence suggests that some methodologically sound evaluation plans are never carried out.

This relative lack of quality evaluation research has created difficulties for the national drug court movement and local programs alike. At various times, scholars and politicians have questioned both the effectiveness and efficacy of drug court programs and some continue to argue that drug courts are not worth the money being spent. While local drug court programs continue to build support and thrive, the federal resource allocation to drug courts is open to challenge and the movement has limited solid evaluation research to refute its critics. Further, no less than 35 states have appropriated funding for drug court programs. While some of these appropriations are pass-through funds from federal programs, others are direct legislative appropriations from general funds, and many are a combination of the two. A recent calculation of state appropriations for drug courts totals close to \$150 million annually (Huddleston et al., 2005). To some degree, all of these appropriations are dependent upon the ability of local programs to document their work and report their outcomes.

### **A National Strategy**

The purpose of this paper, therefore, is to promote quality research at all levels for drug courts by providing a uniform and manageable process evaluation strategy for local programs (see the other articles in this issue for information on recidivism and performance measurement). These methods can be used across the spectrum of drug court programs to allow local jurisdictions to answer questions from stakeholders and funding agencies, as well as promote sound management practices at the local court level. It is clear that research practices can be improved by providing a uniform baseline for evaluation and measurement. This paper

concentrates primarily on *adult* drug courts. While many of the ideas can easily translate to other problem solving courts, the scope of this article is limited to promote research accuracy. Adult drug courts are an appropriate focus mainly due to their prevalence; NDCI places the number of adult drug courts in 2004 at 811 nationally (Huddleston et al., 2005).

It should also be noted that the methods suggested in this paper are not the only ways to gather useful information regarding drug court processes and performance—local court programs need to direct their own research to benefit their own programs. Furthermore, it should be understood that the quality of research depends heavily on access to data and the availability of resources. With this in mind, some additional suggestions for improving research design beyond the baseline requirements are found in the following pages.

This paper focuses on six important questions related to drug court evaluation. The first two questions deal with the general elements of process evaluations and the issue of who should conduct these evaluations. The third and fourth questions relate to the actual performance of local drug court program evaluation, including a discussion of the critical elements of conducting this type of evaluation as well as what data are needed for the purpose of answering these questions. The fifth and sixth questions relate to the issues of methodological rigor and evaluation design.

## **WHAT ARE PROCESS EVALUATIONS?**

[8] Generally speaking, evaluation research refers to a purpose rather than a specific methodology (Maxfield & Babbie, 2005). Simply put, process evaluations should be tools for managers and stakeholders as they seek to maintain successful programs, enhance services, and promote research-based best practices within programs (Rossi & Freeman, 1989). Evaluation research is a means by which programs

can be opened up to determine the extent that they are achieving program goals and managing their activities in an effective and efficient manner. Likewise, the intended audience must motivate process evaluation. In many cases, certain audience members (i.e., state administrators) will ask specific questions that evaluators can address in their reports. Two primary questions drive process evaluation: policy implementation and the achievement of program goals (Maxfield & Babbie, 2005).

Further, evaluation research is commonly defined using three important constructs. The first is the use of a systematic approach in synthesizing evaluation plans (Rossi & Freeman, 1989). That is, the plan must be designed and implemented in a strategic, careful, and consistent manner. This is particularly important since the evaluation often relies on retrospective data, which may or may not have been collected and managed in an easily accessible format. This systematic approach applies not only to the collection of information, but also to the second major construct, the critical analysis of information (Rossi & Freeman, 1989). It is clear that simply collecting information is insufficient to constitute a valid evaluation; it is critical that the information be carefully *analyzed* by individuals or teams that understand the underlying principles guiding program practice. Finally, evaluation research must provide *useful* feedback. Evaluative feedback that is difficult to understand or meaningless to the consumers serves little purpose for the program, and thus is not practically useful, though perhaps highly advanced and descriptive (Bachman & Schutt, 2003).

## Defining Terms

With this in mind, it is important to discuss some basic evaluation research terminology. Evaluators generally consider four *terms of art* when developing evaluation methodologies (Bachman & Schutt, 2003). The first of these terms is *inputs*. Inputs can be considered to be any of the raw



materials that enter the program. For example, drug court inputs tend to be clients, program staff, and additional resources. The second term, *program process*, refers to the treatment and/or services provided for clients in the program, as well as the policies that guide the delivery of those services. Drug courts rely on a variety of process mechanisms to create positive effects including sanctions and incentives, substance abuse treatment, and ancillary services. Third, *outputs* are short-term products produced by the program process. For drug courts, this term could be used to refer to the number of hours of substance abuse treatment received by a client or the number of urine screens that a client provides. Finally, *outcomes* are the impacts that the program has on its participants (Bachman & Schutt, 2003). Often the terms *outcome evaluation* and *impact evaluation* are used interchangeably. However, there is at least one subtle difference between the two: Impact evaluations tend to focus on large scale measures of quality of life beyond the particular client, while outcome evaluations tend to focus on the effects of the program or policy on a particular participant.

## Measurement

Process evaluators must consider each of these evaluation terms (or constructs) and find appropriate means for measuring them within the context of the program. Measures that are too broad or too narrow often tend to over or underestimate program effects. It is critical, therefore, that evaluators carefully consider the variables used to measure client background and risk, program activity, outputs, and outcomes. The best means for identifying the important variables is through a process that first defines the questions and then links the questions to variables that can be sufficiently applied based on existing research or a researchable hypothesis. Recent reviews of drug court evaluation plans by the NDCI research team revealed some significant uncertainty about how best to measure the

important evaluation questions facing drug courts. Often the questions were appropriate but the means identified for answering them were unexplained, unclear, or inappropriate. For example, many evaluation plans did not consider previous treatment failures as an important client characteristic for evaluating their target population. Many simply considered legal measures (i.e., criminal history) as the only defining social variable apart from simple demographics. Existing research, however, clearly shows that certain drug court models and activities perform better with particular types of substance abusing offenders than do others (Marlowe, Festinger, & Lee, 2004).

*Operationalization* is the term used by researchers to define the process of making a construct measurable or turning concepts into variables. It is the act of taking a term like “recidivism” and making it measurable and comparable. It requires an understanding of the definition of the term (i.e., offender committing an additional criminal act after being arrested, charged, or convicted for a criminal act) and creating a meaningful way to measure it (e.g., arrest on a new charge). Often the operationalization of a construct is not exact, as demonstrated by the example above. However, it is important that the method used for measuring the construct be theoretically defensible given the context. The recently completed monograph by NDCI’s National Research Advisory Committee, *Local Drug Court Research: Navigating Performance Measures and Process Evaluations*, suggests using arrest data for analysis of recidivism (Heck, in press). Clearly, arrest data have weaknesses as measures of actual criminality but, given the theoretical defensibility of the choice, constraints of data collection, and the length of time required to get through general court proceedings, it was decided by the committee that this measure would be the best for the purposes of performance measurement.

## **Local Drug Court Process Evaluations**

Drug court process evaluations are tools to be used by programs for improvement and should provide interested parties with a glimpse into the workings of a drug court program—specifically, it should elucidate how the operations of the court produce its effect. These evaluations are focused upon the *how* and *why* of drug court activity. Minimally, a process evaluation should include fundamental descriptive statistics (e.g., simple summaries of certain samples or measures such as the number of men and women in the program, the number of court appearances, the modality of treatment offered) and use these to answer questions concerning the level to which programs are meeting their operational and administrative goals. One common process question focuses on the extent to which the local program is reaching the population it was chartered to serve. By definition, drug courts target particular types of offenders. Those eligible often include offenders with no prior violent history and substantial addiction problems. After reviewing the program's target population goals and comparing this to the type of offender the program is actually accepting, an evaluator may, for example, suggest the refinement of the program's target population, or a refinement in the decision-making process used to accept certain clients in order to make better use of limited resources. By focusing on evaluating the target population, a court is able to better understand its own screening process, as well as evaluate the suitability of its ideal client group, given the resources available.

A comprehensive evaluation of a drug court program should take into consideration the structure and process of the program in addition to examining program impacts (e.g., participant outcomes). Ideally, evaluations should examine how program structure and process contribute to found impacts. In this way, evaluators can help programs answer not only *whether* the drug court works, but also *how* the drug court works. Moreover, evaluators can help program

administrators understand how the internal functions of the court affect client behaviors during their time in the program and beyond (Goldkamp, White, & Robinson, 2001; Longshore et al., 2001). Due to several factors, including a lack of resources, drug court evaluations often only report on client outcomes and do not include a description, discussion, or analysis of important and contributing program process elements. Further, these reports often focus upon outcomes and measures that are unrelated to the program's goals. Given that drug courts operate as a function of local interests, needs, and resources, the lack of attention to (or at least presentation of) court process elements in these very localized programs has also hindered the drug court movement as a whole, as well as the development of a broad and comprehensive process evaluation model that could help guide programs and their evaluation efforts. Though drug court professionals and practitioners have recognized this as an important function to the future sustainability of the drug court movement, the lack of education, training, and technical assistance on the mechanics of conducting an appropriate and methodologically sound process evaluation has often stalled this undertaking (Office of Justice Programs, 1998). This paper provides a resource for addressing this issue.

## **WHO SHOULD CONDUCT THESE EVALUATIONS?**

[9] Process evaluations must be conducted by objective outsiders with knowledge specific to the area in question. Independent evaluators are less susceptible to political and personal pressures while conducting evaluations. Though many programs use self-evaluation models and have program employees conduct their own evaluations, the multiple purposes of program evaluation are not well-served by these methods. As such, we recommend two primary considerations when identifying a prospective evaluator. First, the evaluator must understand evaluation. There is a widespread assumption that anyone with a higher education degree has a fundamental understanding of program

evaluation. This is simply not true. Some disciplines and educational programs focus on evaluation research while others do not. Further, an understanding of scientific methods, though helpful, does not in itself imply the mastery of the skills needed to actually conduct an evaluation. Program evaluation is a specific skill. While many of the evaluation plans reviewed by NDCI had quality researchers guiding them, the proposed design was often inappropriate for drug court evaluation.

One way to ensure that potential evaluators have the requisite skill-set is to look carefully at their research background. Drug court program managers should ask potential evaluators to share previous research project reports, and inquire about the methods that they would employ for this type of research. While experimental designs provide for excellent research, they are generally not appropriate or necessary for local drug court program evaluations. Program evaluators must be skilled at researching the program as it operates in the real world. This means that evaluators must be cognizant of context and program limitations given the social and political environments in which they exist.

Second, program evaluators should have some substantive knowledge of the fields in which the program operates. For drug courts, this means that evaluators with experience in substance abuse treatment, corrections, and court processes are preferable to those without, or with a background in just one of these areas. And, while it is theoretically possible for evaluators to become familiar with these substantive areas while working on the project, bringing the evaluator “up to speed” in these areas would involve a great deal of time and energy that might be better expended elsewhere. Additionally, the academic knowledge of subjects such as behaviorism and substance abuse treatment (core elements of drug courts) is not easily gained in the short term.

While self-evaluations can be useful for program management purposes, these types of evaluations rarely yield the comprehensive and rigorous insights that can be articulated by a trained evaluator. Many such “self-evaluation” models are simply means by which employees collect their thoughts concerning the program in a uniform way. It would, of course, be incorrect to suggest that this modality could result in no possible positive results. However, as mentioned above, many consider employees who have a stake in their own programs to be less objective when it comes to the concerns of their programs. Simply put, outside stakeholders will often consider these evaluations to be less credible than those conducted by an outside, objective researcher. Importantly, it is often the case that program employees “cannot see the forest for the trees,” and thus lack the ability to view the program in context. Moreover, program evaluators with pertinent experience can often provide ideas and strategies for improving the effectiveness of programs that often go unconsidered by program staff.

Finally, it is strongly suggested that program managers consider the evaluator’s willingness to listen to staff concerns during the selection process. Regardless of the amount of programmatic experience the researcher brings to the discussion, managers must remember that evaluations can be guided to answer specific questions that might plague programs. Evaluators should provide systematic analysis of all of the aspects of drug court program operations. Beyond the basic elements inherent to the drug court process, program managers must direct evaluators to consider questions specific to their jurisdiction. For example, if the process of acquiring new, appropriate clients moves at a pace slower than is optimal, the program manager might share his or her concerns with the evaluator and request particular attention be paid to the topic. In general, process evaluations should provide managers with useful feedback regarding the form and function of their programs, with the intent that this

information guide appropriate program improvements, as well as help to document program quality.

To do this, it is important that process evaluators have a solid understanding of the academic research related to drug courts, addiction, and treatment, and be willing to listen. Process evaluations should be conducted with substantial consideration given to the environment in which a drug court program operates, including the actual day-to-day operations of the court, as well as the theoretical constructs associated with the growing body of literature surrounding substance abuse treatment. Sound process evaluations should provide information that is not only based in the research literature, but that is also practical and locally relevant.

## **WHAT ARE THE CRITICAL ELEMENTS OF PROCESS EVALUATION?**

[10] The drug court model has been well defined since the Office of Justice Programs (OJP) and the National Association of Drug Court Professionals (NADCP) released the monograph entitled “Defining Drug Courts: The Key Components” (NADCP, 1997). In fact, many states have copied the key components in their enacting legislation. This model has been utilized with success for over 17 years and while there remain some questions about what parts of the model are most effective (see Goldkamp et al., 2001), there is considerable evidence to suggest that the totality of the approach is effective for retaining clients in treatment and promoting positive outcomes. Thus, one of the important evaluation questions must focus on the integrity of the model as applied by the particular program. These key components are not difficult to operationalize and as such should be included in program evaluations. While it may be the case that variations from this model have developed over time that provide improved services for drug court clients, these variations themselves can be valuable lessons that should be discussed in the analysis of the active program.

At a minimum, there are some basic elements that should be considered in any systematic process evaluation of drug courts. By definition, the following elements should be common to *all* drug court programs:

1. **Program Goals** – Drug court evaluators should examine the extent to which programs are meeting their stated or written goals. Suggestions should refer to meeting these goals more successfully or, alternatively, changing the goals to be more practical or relevant. As program goals often are broadly stated, it is incumbent upon the researcher to define these goals in a manner that is meaningful to program management. Many states have specific program goals as part of enacting legislation for drug courts, and it may be important to review these larger goals as part of the evaluation project. Supplemental to this is a determination of whether the program is operating as designed, particularly since it is not unusual to find that often there is a discrepancy between how a program was implemented versus how it was *intended* to be implemented (Longshore et al., 2001).
2. **Target Population** – It often is difficult to specifically define the population of offenders that a drug court program serves, considering the eligibility requirements that may or may not relate to the suitability of the client. However, it is essential to the operation of drug court programs that they be able to concisely identify the population they hope to serve and determine the extent to which they are reaching that intended group. Drug court evaluators should examine drug court client intake in terms of the program's stated goals (court goals as well as legislative, if applicable), resource limitations, and the universe of those who could be eligible for the program. A common complaint among many drug court programs is the inability to stay at full operating capacity. This problem can be researched and suggestions made through



a thorough analysis of client intake procedures and target population goals.

3. **Substance Abuse Treatment** – One of the aspects of drug courts that separates them from nearly all other justice system interventions is substance abuse treatment overseen using judicial monitoring and enforced supervision. To address this issue, it is crucial that baseline measures of addiction be considered. Evaluators should compare treatment plans with the actual implementation by the court. When possible, it is also important to determine the appropriateness of specific treatment modalities for particular clients. While it is not the purpose of this paper to recommend specific screening instruments, it is important that drug courts document client use prior to the program to enable accurate comparison throughout the program; as such, any instrument used must contain measures of past and present prevalence and incidence of drug use, addiction severity, and drugs of choice. Screening and assessment instruments should contain measures of the appropriateness of particular modalities for particular clients (e.g., American Society for Addiction Medicine criteria) and must be both reliable and valid. Baseline data should then be compared to one or more reassessments of clients' addiction severity, both during and at the conclusion of the program.
4. **Court Processes** – All of the activities of the drug court program should be documented. Researchers should examine graduation, phase advancement, sanctions and incentives, supervision, and the various ramifications of drug testing, as well as the relationship between client needs and services rendered. Behavioral research supports the notion that the magnitude of the sanction or incentive should be proportionally consistent with the precipitating incident, so sanctions and incentives should be measured in relation to client behaviors (Skinner,

1950)). Therefore, it is both possible and desirable to create a ratio of behaviors to sanctions or incentives with the goal of a one-to-one ratio.

A great deal of information has been published recently about behavioral controls of client behavior. For example, Douglas Marlowe and Kimberly Kirby published an article entitled “Effective Use of Sanctions in Drug Courts: Lessons from Behavioral Research” (1999). In this article, the authors describe the need for an overall individualized behavioral plan for clients based upon their personal histories and stations in life. Further, one of the most supported behavioral principles is the idea that certainty (i.e., the likelihood that an action, good or bad, will elicit a response) is perhaps the most important factor in creating client behavioral responses. Thus, measuring the relationship of client behaviors to programmatic responses is critical. Both the perceived magnitude of incentives or sanctions and the application schedule should be reviewed.

Other aspects of court process that bear mentioning are the supervision of clients and the coordination of court activities. Client supervision is one of the key components of the drug court model. Elements of supervision include client contacts and oversight of client activities (e.g., employer contacts). Further, the coordination of service application falls under the rubric of court processes. Questions should be asked about information sharing and team involvement in the decision-making process.

5. **Units of Service** – Drug court clients generally receive a variety of services while in the program. Each of these services should be documented in a manner that helps the program consider the benefits of particular services. A solid process evaluation will report if clients are gaining from particular programs or interventions. A unit of

service is a simple way of measuring and documenting all of the services provided by drug court programs. Included in this documentation should be medical and psychological services, job training and placement services, educational services, and any other service to which the client was linked by program staff. When considering units of service, it is important to document both the referrals and participation in the service provided. Benefits of services would most likely be assessed by asking clients in a consistent manner (i.e., customer satisfaction index) their feelings about the services provided. This attitudinal measurement strategy can provide a useful resource for management.

6. **Team Member Cooperation** – Drug courts are collaborative efforts. Their success or failure is dependent upon the constant “give and take” that replaces the traditional adversarial system. Some method of qualitative organizational research is useful to determine how well the drug court team functions as a unit. One simple method for collecting this type of information involves questioning team members individually as to their perceptions regarding the extent to which their input is considered when decisions are made by the drug court team.
7. **Community Support** – Community support is vital to program success. In some jurisdictions, the voting community selects team members, and courts often use local businesses to provide token incentives. There clearly is value to program management exploring the reactions—either positive or negative—to the drug court in the community it serves, as the court may eventually need local funding and support to survive. With this in mind, it is often valuable to assess the support of stakeholders and community leaders. This can be done using a survey or questionnaire asking specific questions

about their understanding of the model and its implementation.

### **Drug Court Planning Process**

During the course of conducting the process evaluation, particularly the gathering of information through interviews with team members and other important stakeholders, an ancillary component should involve the review of the drug court's *planning* process. Understanding how and under what conditions the court came into existence can help inform an understanding of current processes and protocols, especially if a program has evolved since its inception in response to particular, and often unanticipated, circumstances or resource constraints. An understanding of this important component of the process evaluation can be gauged by asking questions such as:

- Were all appropriate key players brought in to serve as part of the drug court team (to help develop goals, objectives, policies and procedures, and the mission statement)? If not, who was missing? Was a representative from the mental health community at the table?
- Were enough team members assembled?
- Were adequate and appropriate planning trainings offered to all team members?
- Were all of the available community resources documented?
- Did all team members sign a release of information in order to share confidential information with each other? Were team members trained on confidentiality?
- How was the target population defined?

Beyond these seminal questions, drug courts must be considered as organisms that are growing and redefining themselves on a regular basis (Carey & Finigan, 2004).

There is often attrition among the ranks of drug court practitioners working in a particular program. Further, advances in treatment modalities and changes in behavioral approaches force programs to be somewhat flexible in their activities. And despite the fact that there is often natural incongruity over time in program operations, evaluators must consider these factors as part of the growth of the program rather than as separate incarnations. These factors require that a few additional questions be asked of the program:

- What programmatic or personnel changes have occurred over time?
- Are on-going training opportunities (including in-service and cross-trainings) provided to and utilized by all team members? Are these trainings worthwhile?
- What data collection system is being used? Or, how are records kept? Is the current system (computer or paper) working well for all team members?

### **WHAT DATA ARE NEEDED TO COMPLETE THESE EVALUATIONS?**

[11] To answer the research questions mentioned above, there is a significant amount and type of data required. For this reason, trainings supported by BJA require that an evaluator be part of the initial planning team. The purpose of the involvement of an evaluator is to ensure that goals and objectives are measurable and meaningful, and to assist in the effort to collect the appropriate data. History suggests that this is one area in which drug court programs have failed to help themselves. As drug courts are local collaborations of disparate actors, often the programs rely on each of the partners to collect their own data in their own manner while important coordination and operational data is left behind. Indeed, NDCI's review of evaluation plans suggests that the single largest problem facing drug court evaluators is the lack of good data in useable form.

While a comprehensive list of data elements for drug court process evaluation and performance measurement is provided in the forthcoming research monograph (Heck, in press), there are some things that bear mentioning in this article. The first is that data must be collected on a consistent basis that provides for *reliability* and *validity*. Reliability means that the concepts and variables are measured consistently over time (Senese, 1997). This is a significant problem in drug courts, especially given the fact that programs have been buffeted with a variety of data collection scenarios over the course of their existence. Often, programs change midstream to adopt the newest model for data collection. While this may be in the best interest of the drug court for future data collection, it is the responsibility of the staff to ensure that existing data is not lost. This becomes a serious difficulty for researchers and evaluators as they attempt to track and document the historical activities of the court programs in relation to the outcomes.

Validity refers to the extent to which the data accurately reflect the operationalization of the concept or variable. There are four types of validity often considered when making judgments about data. *Face validity* is an assessment of the validity of the data based upon “what makes logical sense” (Senese, 1997). That is, is it reasonable to assume that the measures used accurately depict the construct being measured? For example, it makes some sense that arrest data would be a good measure of criminal activity and therefore a valid measure of recidivism. The second type of validity is *predictive validity*. Predictive validity refers to the extent to which the data accurately predicts the concept. For instance, appropriately applied sanctions and incentives lead to improved client behavior. Thus, a measure of the temporal proximity (celerity) between the action and the court response provides a good measure of the implementation of the behavioral model. Third, there is *content validity*. Content validity requires multiple measures of the same effect. Going back to our example of sanctions

and incentives, it would be important to consider swiftness, certainty, and appropriateness of the sanction or incentive to measure the implementation of the behavioral model. Insofar as these variables work together to predict the outcome, they exhibit a high degree of content validity. Finally, there is *construct validity*. Construct validity refers to the extent to which the measures reflect what is theoretically predicted by the research design. In drug court theory, it is assumed that the confluence of the behavioral, supervisory, and treatment processes lead to client success. Thus, measuring the behaviorism (i.e., incentives and sanctions) in the program process would have construct validity as part of the overall model (Senese, 1997).

Thus, generally speaking, the data elements that are collected by drug court programs must be both reliable and valid. The second important point is that drug court data collection must also capture the important variable of time. Time is generally captured by date stamping all drug court activities. For example, it is important to collect information regarding the date in which a drug court infraction occurred as well as the date when the sanction was applied. These dates allow evaluators to measure the time gap to effectively consider the issue of celerity. It is impossible to measure client performance and improvement without documenting the dates of all activities. This should perhaps be understood intuitively, but unfortunately there are many examples of data collection efforts, particularly in areas of drug testing and incentives and sanctions, which forget this important component.

The third major point is that the best time to start collecting this data in a uniform manner is now. The lack of valid and reliable data from which to assess drug court program performance creates a series of issues for evaluators and researchers. The best way to address these potential problems is to avoid them in the first place. This can be done by carefully conceptualizing the model using the available

resources and working with an evaluator to ensure that the correct data are being collected in a usable format from the inception of the program. If the program is already operational it is strongly recommended that these data concerns be addressed as soon as possible.

Several states and localities are developing or have developed standardized and comprehensive electronic systems for capturing the important data. It is highly recommended that these systems serve as more than just data repositories. There are myriad examples of social programs that have developed data collection systems that have been poorly managed and provide little service to the local programs. Indeed, it seems that this lesson is generally hard to learn, as data collection systems for drug courts are still being developed even at the national level. The technology exists to create case management systems that provide the users with utility and promote the input of valid data for the purposes of managing local and state programs. As more drug court specific systems are developed, the costs are shrinking and the benefits of such a system are tremendous both for program evaluation and in a broader sense, sustainability.

Each of the basic elements mentioned above have attendant variables associated with them. Using the concepts of reliability and validity as a guide, evaluators, program managers, and stakeholders need to develop a data collection enterprise that accurately reflects the program components based upon the theories guiding the model. Again, expediency is critical. It is not enough to allow the partners in the program to collect their own required data, as the important elements of cross-pollination will be missed and the likelihood of identifying gaps in services will be reduced.

It is also important to remember that drug court programs involve treatment providers that collect confidential information that is subject to the Health Insurance Portability



and Accountability Act (HIPAA, 1996) and federal medical confidentiality regulations (CFR 42). These standards require careful management of data and must be followed to ensure continued licensure. There is some confusion about requirements related to sharing health care information. Many people incorrectly assume that HIPAA disallows the sharing of any health care information, but this is not strictly accurate. HIPAA does not limit the ability of direct care providers to share information as long as it is for the benefit of their patients' treatment. There are, however, substantial requirements related to informing and gaining consent from patients (U.S. Department of Health and Human Services, 2003). Evaluators must work with program management to develop the appropriate waivers for clients and to make sure that confidentiality lines are not crossed.

### **WHAT DOES IT MEAN TO BE “METHODOLOGICALLY RIGOROUS?”**

[12] Methodological rigor refers to using empirical and scientific models for analyzing cause and effect. While strong in developing correlations and relationships, social science as a whole has difficulty establishing true causality using the scientific method, and drug court research is no exception. While there are ongoing studies that use random sampling and empirical design (see Marlowe et al., 2004), most local courts and evaluators do not have the available resources to conduct such research. Indeed, by design, evaluation research is not meant to create high statistical significance and be able to claim true scientific causality.

That is not to say, however, that evaluations should not be rigorous in following accepted protocols and methodology. In fact, one of the major problems with many current evaluation endeavors is that they do not follow any uniform protocols. Professional evaluators have developed general standards which are appropriate to apply to drug court evaluation. These standards include *Utility Standards*,

designed to ensure that user's needs will be met; *Feasibility Standards*, designed to ensure that the evaluation will be "realistic, prudent, diplomatic, and frugal"; *Propriety Standards*, designed to ensure that evaluations will be legal and ethical; and *Accuracy Standards*, designed to ensure that the measurement obtained matches the actual value of the variable being measured (Sanders & Joint Committee for Standards for Educational Evaluation, 1994). These standards can be referenced and used as tools for guiding contracts with evaluators as well as for assuring the quality of the evaluation product. While they do not necessarily guarantee academic rigor, they do serve as a good starting point for dialogue with potential evaluators.

Additionally, there are several methodological issues in evaluation that can be avoided by using researchers who understand the foundations of scientific inquiry. It is often the case that evaluators try to do too much with the limited resources available to them. While this effort is somewhat laudable, the result can be poorly conceived research, which ultimately leads to a continued lack of credibility for drug court research as a whole. One common problem relates to attempts to claim the use of experimental design in local program evaluation when there clearly is no true experimentation involved.

## **WHAT ABOUT EXPERIMENTAL DESIGN AND COMPARISON GROUPS?**

[13] In order to assess the relative impacts of drug courts, it is necessary to compare the program effects to those of other similarly situated offenders. However, this comparison is not unilaterally required for process evaluation. Process evaluations can serve the purposes of documenting program development, assessing the extent to which goals and objectives are being met, and ensuring fidelity to the model, without comparison groups. Many times this is enough for program managers and stakeholders. Indeed,

doing this much in a thorough and useful manner is preferable to ill-advised attempts to create a comparison.

However, comparison is the only way to document the impact of the program on clients relative to other interventions. True experimental design is the gold standard for this type of research. Using random assignment of subjects and controlling for extraneous pressures provides researchers with the best method for claiming causality (Campbell & Stanley, 1963). Random subject assignment allows researchers to argue that the groups in the study are equal. This “equivalence” is central to making claims about the true effects of the intervention. Unfortunately, experimental design is extremely difficult to perform in a real-world, criminal justice setting and often inspires questions concerning ethics and fairness. These concerns generally focus on issues of equal access when considering designs that exclude groups from accessing drug court services. Additionally, local programs often do not have the resources (both financial and in the number of available clients) to perform such studies.

The next best approach for documenting the effects of programs is undertaking a quasi-experimental approach using *comparison groups*. Comparison groups do not afford the researcher the ability to claim “equivalence” between the treatment group and the non-treatment group, however, they do, in more general terms, provide evaluators with some measure of program effect. Comparison groups are only as good as the specificity upon which they are being compared. For example, comparing drug court clients to all offenders in a state is a weak comparison. And, while weak comparisons can sometimes be better than no comparison at all, there are many factors that make these two groups systematically dissimilar and thus render claims of relative program success quite tenuous.

In developing comparison groups, researchers should first identify the primary and secondary factors that are related to outcomes as suggested in the literature and match their comparison groups on these items. For drug courts, these factors include some basic demographics (e.g., age and gender) as well as research-driven individual factors (i.e., anti-social personality disorders, criminal history, and previous treatment failures). Some of these variables are continuous (they can be scaled), while others are dichotomous. Thus, it is important to match continuous with continuous and dichotomous with dichotomous. Often, the availability of data determines the extent to which comparisons can be made. Great care must be taken in the selection of these groups (Maxfield & Babbie, 2005).

A common mistake made by drug court programs has been to use dropouts and program failures for comparison. There are at least two reasons why this is inadvisable. First and foremost, there are clear systematic differences between program completers (or graduates) and those who leave the program before completion; these differences are impossible to disaggregate during analysis. Many hypotheses exist about why some individuals are more likely than others to complete drug court programs. Almost all of these hypotheses are untested and unproven and as such, statistical control for these variables sheds little light on the variation. Second, while many program failures and dropouts leave in the early stages of the program, it is impossible to identify the relative effects of limited exposure to the program and thus the analysis will be clouded by uncertainty. Clients who stay in the program for any period of time are exposed to the model and thus are tainted for true comparison.

## **CONCLUSION**

The eminent philosopher Thomas Kuhn suggests that science is designed to solve puzzles (1970). Program evaluation is an important tool for program management and

improvement, as well as for solving puzzles such as what makes these drug court programs work. It provides a means by which drug courts can document progress, memorialize actions, and maintain accountability as well as the fidelity of the model. While there have been many excellent program evaluations conducted on drug court programs, there remains considerable confusion and inconsistency surrounding this important endeavor. It is hoped that this article will help to create a foundation upon which drug courts can build an evaluation model that leads the field of criminal justice and supports continued growth and improvement.

---

## REFERENCES

- Bachman, R., & Schutt, R.K. (2003). *The practice of research in criminology and criminal justice* (2<sup>nd</sup> ed.). Thousand Oaks, CA: Pine Forge Press.
- Campbell, D.T. & Stanley, J.C. (1963). *Experimental and quasi-experimental designs for research*. Boston: Houghton Mifflin Company.
- Carey, S.M. & Finigan, M.W. (2004). *Adult drug court typology interview guide*. Portland, OR: NPC Research, Inc.
- Code of Federal Regulations Title 42.
- Goldkamp, J. S., White, M.D. & Robinson, J.B. (2001). Do drug courts work? Getting inside the drug court black box. *Journal of Drug Issues* 31(1), 27-72.
- Heck, C. (in press). *Local drug court research: Navigating performance measures and process evaluations*. Alexandria, VA: National Drug Court Institute, National Association of Drug Court Professionals.
- Health Insurance Portability and Accountability Act (HIPAA). (1996, August 21). Public Law 104-191. 104<sup>th</sup> Congress.
- Huddleston, C.W., Freeman-Wilson, K., Marlowe, D.B., & Roussell, A.P. (2005, May). *Painting the Current Picture: A National Report Card on Drug Courts and Other Problem Solving Courts, I*(2). Alexandria, VA: National Drug Court Institute, National Association of Drug Court Professionals.
- Kuhn, T.S. (1970). *The structure of scientific revolutions* (2<sup>nd</sup> ed.). Chicago: University of Chicago Press.

- Longshore, D., Turner, S., Wenzel, S., Morral, A., Harrell, A., McBride, D., Deschenes, E., & Iguchi, M. (2001). Drug courts: A conceptual framework. *Journal of Drug Issues* 31(1), 7-26.
- Marlowe, D.B., DeMatteo, D.S., & Festinger, D.S. (2003, October). A sober assessment of drug courts. *Federal Sentencing Reporter*, 16(1), 113-128.
- Marlowe, D.B., Festinger, D.S., & Lee, P.A. (2004). The judge is a key component of drug court. *Drug Court Review* IV(2), 1-34.
- Marlowe, D.B. & Kirby, K.C. (1999). Effective use of sanctions in drug courts: Lessons from behavior research *Drug Court Review* II(1), 1-31.
- National Association of Drug Court Professionals. (1997, January). Defining drug courts: The key components. Washington, DC: Bureau of Justice Assistance, U.S. Department of Justice.
- Office of Justice Programs. (1998). *Drug court monitoring, evaluation, and management information systems*. Washington, DC: Author, U.S. Department of Justice.
- Rossi, P. H. & Freeman, H.E. (1993). *Evaluation: A systematic approach* (5<sup>th</sup> ed.). Thousand Oaks, CA: Sage Publications.
- Sanders, J., & The Joint Committee on Standards for Educational Evaluation. (1994). *The program evaluation standards* (2<sup>nd</sup> ed.). Thousand Oaks, CA: Sage Publications.
- Senese, J.D. (1997). *Applied research methods in criminal justice*. Chicago: Nelson Hall Publishing.

Skinner, B.F. (1950). *Science and human behavior*. New York: McMillan Publishing.

U.S. Department of Health and Human Services. (2003). *Fact sheet: Protecting the privacy of patients' health information* [Online]. Available: [www.hhs.gov/news/facts/privacy.html](http://www.hhs.gov/news/facts/privacy.html).

U.S. Government Accountability Office. (2005, February). *Adult drug courts: Evidence indicates recidivism reductions and mixed results for other outcomes*. Report to congressional committees. Washington, DC: Author.